

# Tensions within the Ministry of Provenance: Reflections on Co-Creating a Research Game together with Artists

Richard Wetzel<sup>1</sup>, Khaled Bachour<sup>2</sup>, Martin Flintham<sup>3</sup>

## Abstract

*Background.* **Research games** are challenging to design as they seek to fulfil a research agenda as well as work as a game. We have successfully **collaborated with a group of artists** in a research game about people's perception of provenance called The Apocalypse of the Ministry of Provenance (MoP). The web-based game ran for 6 months with a total of

---

<sup>1</sup> Lucerne University of Applied Sciences and Arts, Switzerland

<sup>2</sup> University of Lincoln, United Kingdom

<sup>3</sup> University of Nottingham, United Kingdom

### Corresponding Author:

Richard Wetzel, School of Information Technology, Lucerne University of Applied Sciences and Arts,

Suurstoffi 41b, 6343 Risch-Rotkreuz, Switzerland

Email: richard.wetzel@hslu.ch

1004 players signing up over its lifetime with 490 consenting to their data being used for research purposes. While the game allowed us to answer our provenance-related research questions, in this article we look at the game design process of such a collaborative research game.

*Aim.* The **co-creation approach created tensions** that had to be carefully negotiated between everyone involved. The purpose of this article is to investigate the nature of these tensions, what has caused them, and how we managed (or failed) to mitigate them. This leads to recommendations for future researchers co-creating a research game with artists.

*Method.* We use the form of a **post-mortem reflection** on the development of the game, based on our own experiences, a one-hour long interview with the two artists involved, and post-game phone interviews with players (n=8).

*Results.* We identify the following **three tensions** that had a high impact on the overall process: 1) Translating research questions into engaging gameplay elements; 2) Creation of research-relevant content by artists; 3) Artistic vision conflicting with research agenda. We contextualize these tensions by describing **six vignettes** concerning our collaboration in

**rich detail** that highlight the salient issues of the overall process and resulting game from different perspectives. Lastly, we present **seven mitigation strategies** on how to deal with the tensions or prevent them from arising.

*Conclusions.* A collaboration with artists for the purpose of creating a research game is a rewarding but also challenging process. Overcoming the resulting tensions is possible by utilizing mitigation strategies that need to be implemented jointly between researchers and artists to guarantee the success as an engaging research game.

#### **Keywords:**

artists, co-creation, game design, post-mortem, provenance, research games, serious games

## **Introduction**

Creating a successful serious game is a challenging task as one has to not only create an engaging game but the same game also needs to satisfy the desired serious outcome (Winn, 2009). In research games, this serious outcome is the ability to investigate the subject matter of the research. Games have been used to investigate many different domains, including health and wellbeing (Beaubien & Baker, 2004), disaster response (Haferkamp,

Kraemer, Linehan, & Schembri, 2011), presence (von der Pütten et al., 2012), or to explore new forms of games themselves (Wetzel et al., 2009).

In our case, we were interested in examining the perception of people on provenance.

Provenance is a record of the history of an object or a piece of data that allows to evaluate the validity of the information at hand. Modern examples for provenance are the edit history of Wikipedia articles or how e.g. the wine and meat industry provide a complete chain of how, when, and by whom the product in question was created.

As provenance is not something most people consciously deal with on a daily basis, we decided against a traditional lab study. Instead we deployed our research game *in the wild* (Benford et al., 2013) as we saw naturalistic and long(er)-term engagement crucial for players to immerse themselves more deeply with the concept of provenance. In addition, we invited artist group Urban Angel as partners to co-create the game with us from scratch with our provenance agenda as the starting vision. This was motivated by the observation that collaborating with artists can lead to more engaging, more contextually relevant and more impactful games and experiences (Benford et al., 2013).

The Apocalypse of MoP (AoM) is a web-based alternate reality style game that engages players with provenance over an extended period of time. AoM is set in a dystopian mirror

universe of our own reality where the Ministry of Provenance (MoP) controls the flow of information and thus truth itself. The narrative-heavy game was live for six months during which 1,004 players registered to play.

This article is a post-mortem on our experiences with the co-creation process of AoM. A post-mortem is a popular way in the game industry to reflect on the development process of a published game (Shirinian, 2011). Reflecting on one's own practice is a valuable instrument for improvement (Schön, 1983). Despite their usefulness, in both HCI as well as in serious games research specifically, such retrospective accounts are unfortunately the exception rather than the rule (Khaled & Ingram, 2012). In this article, we provide such a detailed account of our own experiences on how we as researchers collaborated with a group of artists. The goal of our retrospective review is to unpack the co-creation processes of the game and how artistic vision and research agenda at times were seemingly in conflict with each other. Where the artists primarily wanted to create an engaging game, we needed to make sure that the game provided us with enough relevant data to investigate our original research questions regarding provenance. While we believe that our game was ultimately successful in answering these questions, we also think that more can be learned

from our experience about the overall process of a tight collaboration with a group of artists, which in turn is the focus of this post-mortem.

The core of the article is an in-depth look at three tensions that arose during our collaboration: 1) Translating research questions into engaging gameplay elements; 2) Creation of research-relevant content by artists; 3) Artistic vision conflicting with research agenda.

## The Co-Creation Process

Games used for research purposes can be created by the researchers themselves as in the case of Atomic Orchid (Fischer, Jiang, & Moran, 2012), or together with domain experts as seen in Child Interview Simulator (Hart, Iacovides, Adams, Oliveira, & Margoudi, 2017). Collaborations such as the latter are rather common in the field of serious or applied games. Typically, you have game designers and developers as one part of a multidisciplinary team, joined by domain experts of the serious or applied content. These teams then have to closely collaborate in order to reach their non-trivial goal: *“Making a good game is hard. Making a good serious game is even harder. The reason it is so difficult is that rather than*

*simply trying to optimize the entertainment aspect of the game, or the so-called fun factor, one must also optimize to achieve a specific set of serious outcomes.”* (Winn, 2009, p1011)

While there has been substantial work on how to integrate serious content into a game from a game design perspective (Arnab et al., 2015; Catalano, Luccini, & Mortara, 2014; Marfisi-Schottman, George, & Tarpin-Bernard, 2010; Westera, Nadolski, Hummel, & Wopereis, 2008), a look at the actual process of this collaboration and the arising negotiations and tensions has not seen an equal focus. One exploratory study has explored these issues from the view of Dutch game developers with experience with such set-ups (van Roessel & van Mastrigt-Ide, 2011). The authors identified several challenges based on the fact that a domain expert is involved in the process. For the game company it is crucial to *elicit knowledge* from the domain expert in order to be able to integrate the desired content into the game, something described as not being an easy task: *“extracting the relevant and useful knowledge from the domain expert, is an expertise in itself”* (van Roessel & van Mastrigt-Ide, 2011, p. 7). However, as the domain experts are the usual clients of the game company, they naturally hold a certain amount of power over the game – with direct repercussions on the game design. As the domain experts are *“likely to advocate the content part over the game play, chances are that correct content is overemphasized over game play or the entertainment value of the game”* (van Roessel &

van Mastrigt-Ide, 2011, p. 9). Lastly, the authors described that a transfer expert was usually non-existent, i.e. a professional with an educational or pedagogical background to assure that knowledge or skills gained inside the game would then also be applicable outside of it.

Another study highlighted six challenges for games for health derived from interviews with professional game designers (Cheng, Putnam, & Guo, 2016): 1) achieving a balance between player engagement and the game's effectiveness at addressing the health goals; 2) consolidating the different mindsets of game designers and domain experts; 3) measuring the efficacy of games for health; 4) limited resources such as time and budget; 5) achieving sustainable impacts and maintaining lasting player interest; and 6) handling stereotypes associated with gaming. For the purpose of this article, challenges 1) and 2) seem especially relevant as they look at the collaboration process.

Both of these sources highlight relevant issues that we also encountered in our process. However, one significant difference to our work is the focus on the perspective of the game designers. But what are the challenges from the point of view of the domain expert? In AoM, we took on exactly this role (and the role of the game developers) and had to negotiate with the game designers of this project, the artist group.



Researchers have successfully collaborated with artists in the past, e.g. Blast Theory's Can You See Me Now (Benford et al., 2006), Active Ingredient's 'Ere Be Dragons (Davis et al., 2006), or Brendan Walker's Breathless (Benford et al., 2012). Likewise, we had previously worked with Urban Angel on their game Malthusian Paradox (Evans, Flintham, & Martindale, 2014). In all these cases, however, the artists tended to "*typically have a vision, which may be more or less well formed at the start, of what the artwork is to be*" (Benford et al., 2013, p3), and researchers helped to bring this vision to life and then subsequently studied the game being run.

In contrast, in AoM we established a stronger focus on co-creation: Urban Angel were contracted by us and put in charge of the development of the basic premise, the overall narrative, creating all content, runtime management, and engagement with players as well as advertising the game. On our side, we introduced the artists to the concept of provenance, implemented the game, and created tools to support the authoring of content. We also directed the artists to make sure the game would allow us to investigate our research questions, and as such influenced the overall game design. The latter became an iterative process with long discussions as for example the artists would often propose a game element, and we would then together shape it into a research-relevant form. Likewise,

we made suggestions regarding the game design based on our research agenda, which the artists then took on board and transformed into something that would make sense in the context of the game.

Overall, the Apocalypse of MoP therefore is an example of a research game truly co-created by researchers and artists. This allows us to highlight the issues that arose throughout the design, development and staging of the game. Researchers who seek to likewise collaborate with artists will be able to use our mitigation strategies over the course of the co-creation process of research games but also general serious games projects.

## Provenance Research

As AoM, and our research behind it revolves around provenance, we briefly explain the concept and discuss existing HCI-related research around provenance as well as our own interests in regards to AoM.

The W3C defines provenance as *“a record that describes the people, institutions, entities, and activities involved in producing, influencing, or delivering a piece of data or a thing”* (Moreau & Missier, 2013, section 1). As such, provenance is generic information that captures what happened, either in a computer application or in the real world. It offers the

means to verify information and infer its quality, to analyse the processes that led to a thing, and to decide whether or not it can be trusted (Moreau, 2010). It provides the means to attribute aspects of a certain object to its different contributors. The recent emergence of provenance as an important concern in various applications (e.g. establishing accountability, reproducibility and trustworthiness of information) has led to the release of the PROV Standards in 2013 (Groth & Moreau, 2013). PROV is designed to express provenance information in a clearly defined and machine-readable form for heterogeneous environments such as the Web.

Closely related to this PROV standard are studies of provenance concepts in data retrieval with provenance information presented in a variety of visual forms (Del Rio & Da Silva, 2007; Ghorashi & Jensen, 2012) but also with some awareness that users of provenance data may not be database experts themselves (Herschel & Hlawatsch, 2016). As a standard, PROV is beginning to be used, for example to access the version control information resulting from Git (De Nies et al., 2013), or the history of Wikipedia pages as they are edited (Missier & Chen, 2013). However, the usability of the PROV standard itself as a readable representation of provenance remains unstudied, and in general there is little work

on understanding how to adequately visualise different types of provenance, or to understand the requirements of end users (Herschel, Diestelkämper, & Lahmar, 2017).

Provenance concepts – as different from the PROV standard itself – with which people interact daily have been variously directly or indirectly studied by HCI researchers in a number of domains. This has taken the form of augmenting physical objects with technology that enables them to sense or be sensed, affording the ability to attach a digital record to the object, and subsequently building up digital footprints that allow a user to inspect the history or provenance of the object – for example who has owned the Carolan Guitar and where it has been played (Benford et al., 2016), or to hear stories from the previous owner of an item of clothing (Barthel et al., 2013). Being able to reason about provenance is considered important regarding food, e.g. through indicators of country of origin or sustainable production. This has seen the development of conceptual tools that reveal food provenance to support ethical food decision making (Light, 2014). In a very different domain, what the provenance of a news article is and how this provenance is revealed in a social media feed is a current area of research given concerns around so-called fake news (Rubin, Conroy, & Chen, 2015), or investigating how meta-data – arguably

digital provenance data for media – is used in music production (McGarry, Tolmie, Benford, Greenhalgh, & Chamberlain, 2017).

In our own provenance-related research, we were interested in looking at provenance from a user perspective. How can PROV, a standard for storing and transmitting provenance information designed to be machine readable, be displayed so that it is easily understood and digested by humans? How do people's understanding and handling of provenance change over time, especially regarding strategies on how to layout and structure provenance data? Lastly, we were also interested in the *big picture*, i.e. how do people perceive provenance on a personal and community level? What are the ethical and social implications of provenance models in the real world?

## The Apocalypse of MoP

The Apocalypse of MoP is a web-based game where players enter an Orwellian version of our world. A dystopian super-governmental Ministry of Provenance (MoP) monitors and controls every aspect of people's lives. Players join the mysterious underground resistance movement, *Cr0n*, and in turn infiltrate MoP as a double agent.

The game can roughly be divided into two parts that take place on two different websites. Inside the MoP intranet players have to fulfil their daily office job of inspecting provenance graphs and mark any inconsistencies they find in the data. The MoP intranet makes heavy and ironic use of gamification which allows players to gain valuable TRUST points based on their work performance which normally would allow them to access additional work documents, but which they can then abuse to request top-secret documents and leak them to Cr0n by exploiting several backdoors in the MoP intranet. In addition to these gamification elements, the MoP intranet is intentionally designed, according to the vision of our collaborating artists, as an overly bureaucratic behemoth of an organization (Figure 1): a plethora of independent departments with different goals and agendas; propaganda and motivational statements on every page; forms that need to be filled in and signed to request additional forms; and an automated but slightly clunky internal email system. The latter also allows players to engage with Non-Player Characters (NPCs) controlled by the artists: Sandy Spencer is an overly enthusiastic and friendly liaison officer that often gives tips to players while Christopher H. Delendi, the harsh and strict head of the Unit of Auditing and Internal Investigations (or MOPAIN), might grow suspicious of players' motivations and background.

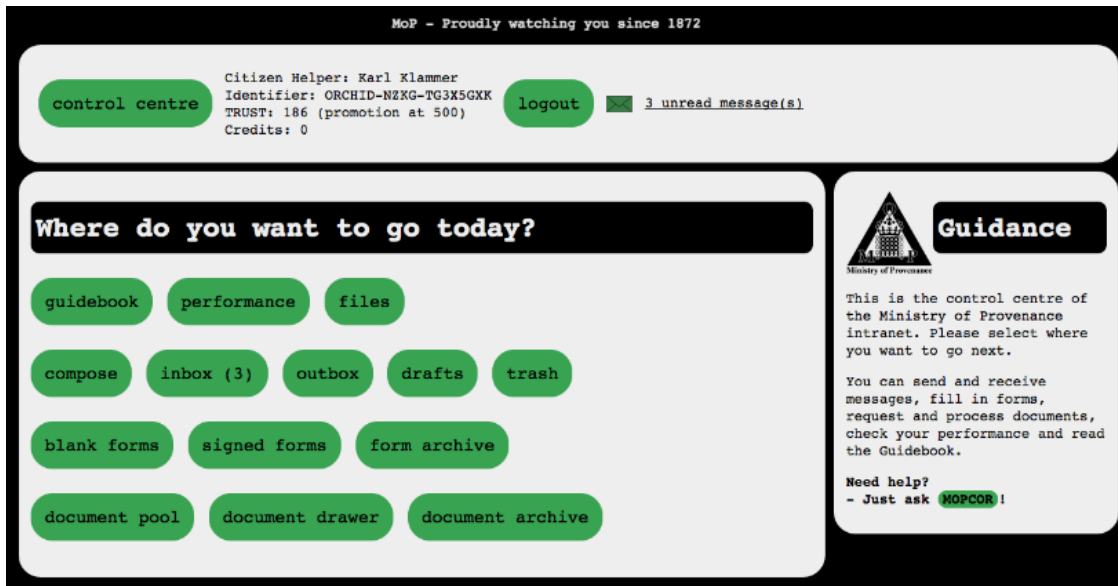


Figure 1. Web interface to navigate the intranet of the Ministry of Provenance.

In contrast, an enigmatic and charismatic leader known only as The Groundsman runs Cr0n. Throughout the game he speaks to the players via short video clips to convey parts of the narrative or pose new questions (Figure 2). Unlike the MoP intranet, players only need to engage with the Cr0n page when a new mission is released. Over the course of 6 months we created a total of 18 missions spread out over four episodes. Each mission starts with a narrative hook where a scene is described between characters that players may or may not have encountered previously in the game. They then get a mission briefing from the Groundsman and need to “steal” one or more documents from the MoP intranet. When they have done so they can open and examine these documents.

The majority of documents were provenance graphs similar to but more complex than the ones players were typically engaging with as part of their daily routine on the MoP intranet – in this instance players had to mark suspicious parts of the graphs that evidenced MoP tampering. Other times players had to solve missions unrelated to provenance, e.g. deciphering codes, looking up GPS locations, or closely observing surveillance footage. In order to complete a mission, players also had to answer questions about it that helped us gauge the players' interpretation of the mission's background. Players then proceeded to a debriefing by the Groundsman and a continuation of the initial narrative which revealed further details about the overarching plot and/or game world. Similar to the MoP intranet, the Cr0n website also allowed players to send messages to NPCs like the Groundsman himself or Agent Dobbin, the Groundsman's second-in-command. This enabled players to receive help when stuck but was also often used to discuss events in the game and influence the overall narrative.



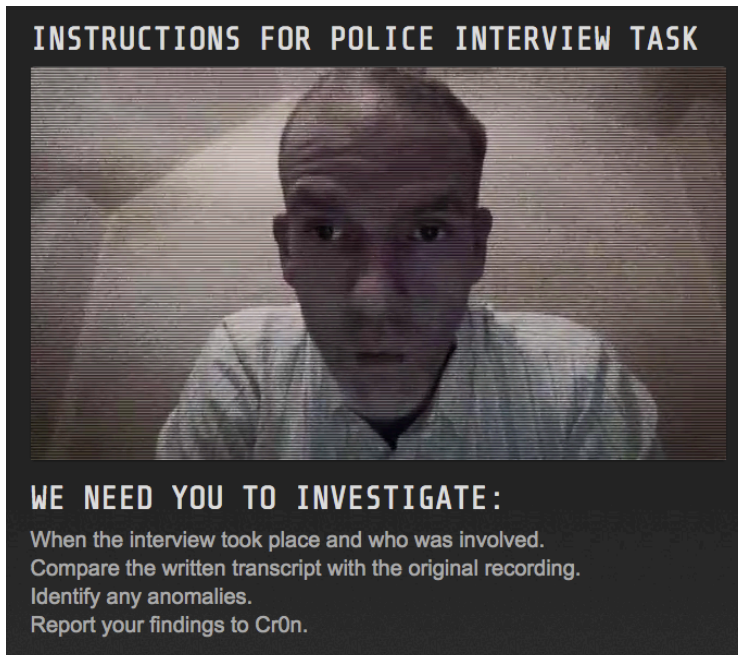


Figure 2. The Groundsman giving video instructions for a new task.

## PROV as Gameplay

Visually, PROV is presented as a graph structure with three different types of elements: entities, activities and agents (Moreau & Missier, 2013). An entity *“is a physical, digital, conceptual, or other kind of thing with some fixed aspects; entities may be real or imaginary”* (section 2.1.1). An activity *“is something that occurs over a period of time and acts upon or with entities; it may include consuming, processing, transforming, modifying, relocating, using, or generating entities”* (section 2.1.1). Finally, an agent *“is something that bears some form of responsibility for an activity taking place, for the existence of an*

*entity, or for another agent's activity*” (section 2.1.3). Between the entities, activities, and agents, there can be a number of different types of relations. The provenance visualizations inside AoM followed these specifications.

Each provenance graph displayed the relationships between the different entities in a graphical way. Players could then rearrange the position of the nodes to make sense of the connections between them. Each node also contained additional data that was displayed on the side of the interface (Figure 3 and Figure 4). The objective of each graph was to find a certain inconsistency in the data presented. Usually, this inconsistency entailed two attributes in two different nodes that have values that do not make sense given the structure of the graph. For example, a PROV graph describing a currency exchange transaction might have the starting sum of money that is an attribute in one node not matching the resulting sum of money in the new currency presented as an attribute in another node. At the MoP side of the game, these attributes simply had to be marked and sent back, with the appropriate forms, to the relevant department for inspection (Figure 3). When correct, TRUST points were awarded, otherwise a penalty was applied.

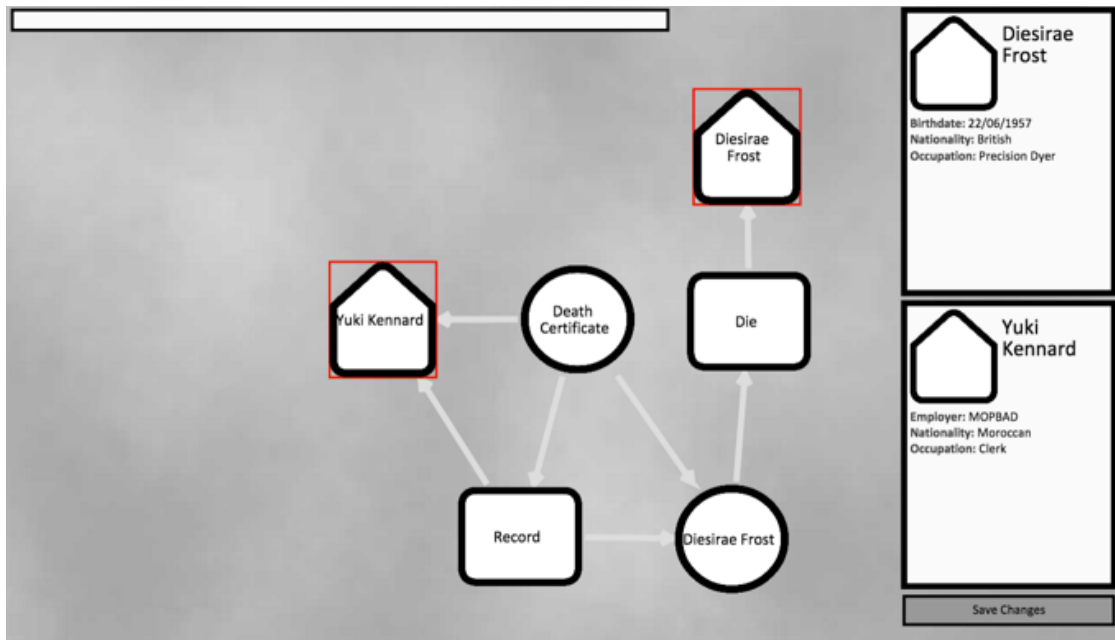


Figure 3. A Provenance grinding task from the intranet of the Ministry of Provenance.

At Cr0n, the graphs were often more complex and contained multimedia elements such as audio recordings and security footage (Figure 4). The players also needed to look for inconsistencies but rather than simply awarding points, the inconsistencies revealed important parts of the narrative by highlighting a fact that the Ministry is trying to cover up, thereby pushing the story forward. The Groundsman would then comment on the revealed inconsistency and his thoughts about future steps for Cr0n. Players then had to wait for the next call to action, i.e. the release of a new mission.

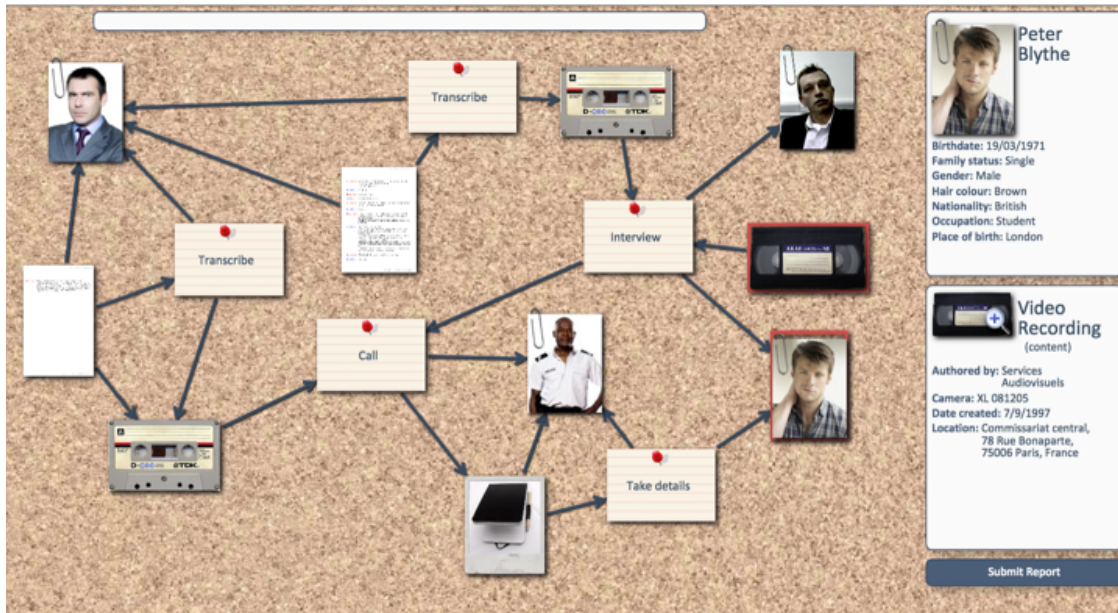


Figure 4. A Provenance story task on the Cr0n website.

## Research outcomes regarding provenance

While the game was running, 1004 players signed up to play. We chose to allow players to participate in the game while opting out of the research; as such, for our provenance-related research we only looked at 490 players who did consent to their data being used for research. Among these, the average age was 28 (SD=10.49), with 334 players identifying as male (68%), 132 as female (27%), and 24 as non-binary (5%).

During the course of the game, we logged all player interactions with the two websites.

Special care was taken to track player interactions with provenance graphs. After the

conclusion of the game, we sent an online questionnaire to all 490 players who had consented to be part of the research, and 41 players submitted responses (25 male, 14 female, 2 non-binary). The questionnaires were a combination of Likert scales and open questions exploring similar issues to the interviews. They were broken up into different sections with each section being relevant only to players who had progressed to a certain level of the game. This allowed us to get feedback from players who reached the end of the game as well as those that quit as early as the tutorial.

In addition, we invited the 40 most active players in the game to conduct an hour-long phone interview. We offered a £10 voucher as compensation for their time. 8 players agreed and were subsequently interviewed (5 male, 3 female). The interviews were semi-structured and were broken up into five phases. First, we explored the player's understanding of and familiarity with provenance followed by a walkthrough where the player was asked to solve a provenance graph using a think-aloud protocol while the interviewer observed the graph manipulation. This was followed by questions about the PROV model itself as well as the in-game interface. The fourth phase explored issues related to history, provenance and the PROV model in a real-world context. Finally, we asked some questions about the game itself and the player's enjoyment of it.

Across both interviewees and questionnaire respondents the progress into the game varied, including both players who stayed with the game for the whole duration of its runtime and others who left after completing the first episode. This allowed us to not only gain insight from our most dedicated players but also the ones we lost along the way shedding light on the various reasons.

Our main findings regarding our original research questions have been published at CHI (Bachour et al., 2015). There we discuss how players perceived the notion of provenance, PROV as a standard, understanding provenance graphs, arranging of provenance graphs, PROV in the real world as well as personal history, privacy, and the ethics of provenance.

Overall, our research allowed us to gain valuable insight on:

- the suitability of the PROV model as a medium of information dissemination for the general public
- the directionality of arrows and the expectation of forward flow
- the social and ethical implications of the pervasive and permanent storage of provenance
- perceptions and attitudes of history and the verifiability of historic records.

As such we see our research as having been successful – however, the game itself also was flawed in a variety of aspects that in turn negatively affected our research. We will discuss

these in the following section with a special emphasis on tensions arising between our research agenda and the agenda of the artists.

## Tensions within The Ministry of Provenance

In this section, we will describe the three tensions between the researchers (us) and the game designers (the artists) that we identified as most salient for our work on AoM. These tensions broadly mirror aforementioned findings and are as follows:

- **Tension 1:** Translating research questions into engaging gameplay elements
- **Tension 2:** Creation of research-relevant content by artists
- **Tension 3:** Artistic vision conflicting with research agenda

We identified these tensions by conducting a post-mortem reflection on The Apocalypse of MoP. Our reflection is based on our own experiences while developing and running the game. Throughout this process we were driven by our provenance related research questions and thus regularly had to consider whether the direction the game was taking was ultimately beneficial for this research. When analysing the data for an academic publication, we likewise noticed flaws in the design of the game that affected our ability to investigate all of our original research questions in the same level of detail.

Our reflection also involves thoughts from the artists and player feedback which allowed us to consider different perspectives of the game (research outcome, artistic vision, player experience).

In order to provide the view of the artists, we conducted a semi-structured post-mortem interview with them that lasted 1h10min. We asked them to reflect on the game and how it affected their practice being a commissioned research game. This included asking them about: the design process and challenges they encountered from wanting to design an engaging game while at the same time also taking the research goals into account; how runtime orchestration of the game was affected by the same constraints; and instances where the requirements of the research content compromised their artistic vision. We also prompted the artists to reflect on specific examples from the overall process that we believed might be exemplary for the aforementioned lenses such as: creating provenance content; requiring players to create their own provenance graphs; including provenance of provenance as a gameplay element; managing direct engagement with players; and the 24h hacking challenge. Quotes are attributed to the two artists Adam and Dominic respectively. In addition, we also used the one-hour long post-game phone interviews with players (n=8) to not only answer provenance related research questions (see previous sections) but also to gauge their overall impression of the game, especially focusing on differences between



artist-driven parts of the game and research-driven ones. If players did not mention certain elements of the game in this part of the interview, we explicitly prompted them about e.g. creating provenance graphs, provenance of provenance, and the 24h hacking challenge. Player quotes are anonymized and attributed as P1, P2, etc.

In the following, we will introduce each of the tensions in the context of AoM. For each tension, we have identified two vignettes of gameplay elements in order to richly illustrate them. The tensions themselves cannot be separated completely from each other as there exists a certain dependency / overlap. Therefore, some vignettes might also apply to other tensions.

### **Tension 1:** Translating research questions into engaging gameplay elements

This tension is perhaps the one that is most generic as it arguably applies to all research games. The research agenda behind the game needs to be embodied by the game itself, i.e. by the gameplay elements: mechanics, narrative, etc. In the case of the Apocalypse of MoP, the grinding tasks, player-created provenance, and provenance of provenance are examples for core gameplay elements derived from the underlying research. As such, their level of engagement is crucial for the overall success of the game as a research method. Despite our best efforts, our game struggled in this regard. While some players found the grinding tasks

engaging, a substantial amount of them found these off-putting. We made similar experiences with player-created provenance and provenance of provenance. With the latter two we also ran into the problem that these could only be accessed after substantially progressing in the game. Due to general problems of player retention, this also meant that only a small fraction of players actually engaged with these parts of the game. In part, this has to do with the linearity of our game. Players were unable to progress to these parts unless they engaged with the grinding tasks. While this made sense from a narrative point, the research questions related to these three gameplay elements did not require this as they were independent enough. In our discussion with the artists it also became clear that they would have tackled the problems quite differently. In fact, these are all examples for instances where the artists would have preferred not to add these as gameplay elements – however they were forced to make the best of it as these were crucial for our research agenda.

### **Vignette 1: Provenance as a grinding task**

The primary game activity that players could undertake at the MoP intranet is probably best described as “grinding”. In games, grinding describes a simple task that players perform repeatedly in order to gain benefits while not necessarily furthering the game narrative. As part of their identity as a citizen helper, the Ministry gave players access to a task pool from

which they could choose provenance graphs to look at and inspect for consistency. Graphs were coming from a variety of different departments (pertaining to traffic, rural affairs etc.). They were presented in five different clearance levels (blue, green, yellow, orange, and red) and a new player would start at the lowest one. In order to gain access to a new level, a player had to process a certain number of graphs correctly. In the narrative back-story the Ministry had introduced gamification approaches to entice their citizen helpers to work harder and more efficiently. There was a total of 15 different templates for the graphs. A script would then continually instantiate the templates filling them with randomized content. This way each graph had the exact same node and relationship structure; however, each graph described a unique occurrence and potentially different inconsistencies.

One of the reasons we implemented the graphs this way was to be able to have a large collection of graph samples that players can work on. During the course of the game, each player would for example engage with a specific template dozens of times. This enabled us to analyse their behaviour over time, examining how the layout arrangement of a graph might evolve and if players became more familiar with its structure. The repetitive nature of these tasks was thus advantageous from a research point of view as it increased the amount of data collected, as well as gave us data about progression of individual user proficiency in PROV.

This however was an issue for the artists as they did not initially appreciate the repetitive nature of these tasks and felt that they would hinder the enjoyment of the game. As such they ensured that some of the provenance graphs players had to deal with in the intranet would be specific to the overall narrative (e.g. confronting players with an orphan re-education scheme that would become relevant later on). The MoP intranet in its entirety was also deliberately set up to enable world building. Without it players would have had something to do only when a new Cr0n mission was released, which was infrequent. Constructing a whole and very bureaucratic online presence of the Ministry also gave MoP a voice untainted by Cr0n propaganda, and consequently helped to give a rich meaning to the grinding tasks.

*Adam (artist): I wouldn't have normally put something like that in an ARG type project that we've done. I wouldn't have people just mindlessly repeat tasks in a way, whereas for the research part, that was a really important thing. And in hindsight, for the game, people, for some reason people enjoy grinding in games, and some people will just keep doing it. [...] It got more and more interesting the more they pursued it. The documents got harder, and you got better situations, and some of them were funny. And I can see why sort of people would wanna see what the next one was.*

Player reaction to the grinding elements of the game was mixed. Some negative responses were received due to usability issues (with some of the flaws being intentional due to the 1984-style bureaucracy). Other players were under the wrong impression as to the level of grinding necessary to proceed in the game, while a relatively minimal amount would have been sufficient.

*P3: It was usually pretty easy. I enjoyed doing it because I like that sort of puzzle. But like I said the only complaint I had was that it would take so long to get the [Cr0n] graphs that you'd need. It got a little bit tedious. Not just the clicking and submitting a whole bunch of forms, but waiting for the graphs to show up was just a little bit tedious.*

Other players appreciated the grinding tasks and even incorporated them into their daily routine.

*P4: The most interesting thing was doing the [MoP] graphs. I enjoyed that. I enjoyed being able to do the graphs every day. I used to have like my little routine where I would get up, do my graphs, go to work, do what I had to do work wise, then come home and I would maybe check if there was more graphs. It was the graphs that made it more enjoyable for me.*

## **Vignette 2: Provenance as a narrative element**

The artists were tasked with creating the theme and the narrative for the game, and to subsequently create appropriate game content and assets. However, this emerged gradually after a series of discussions in which they were introduced to the notion of provenance as a concept, PROV as a technical standard, and conversations around possible dystopian or utopian futures in which PROV and the use of provenance had become ubiquitous. Importantly, these notions were all specific research themes, and we were interested in understanding players' reactions to them.

The setting of the game was inspired by Orwellian 1984—style dystopian governments controlling and monitoring every aspect of life. The artists brought the world alive through videos, such as the Groundsman's briefings or MoP propaganda material, message-based interactions with NPCs, the thematic MoP intranet, and the main narrative and graphs on the Cr0n side. In each episode a narrative bit would give some background to the situation before the Groundsman would call the players into action. Players could also gain a lot of information from interpreting the provenance graphs aside from the inconsistencies that had to be found. The narrative theme of the game arose naturally and in hindsight appeared to have been inevitable.

*Adam: The Ministry of Provenance, that bit came quite naturally. That bit was quite easy, exciting, it seemed like we had a great idea.*

*Dominic (artist): It's the Ministry of Truth, isn't it? It's the whole, sort of, 1984. It's just there, and it's sort of, I think we all got quite excited. That's the sort of stuff we make.*

*Adam: Everything we do is slightly dystopian, techno-y.*

*Dominic: The idea of the Ministry of Provenance and that everything could be tracked is a real possibility. Like what David Cameron said after the shootings in Paris: 'Right we are gonna keep everyone's web browsing history on file for a year'. [...] That sort of stuff exists already. It's a possible real world for ourselves, that could be. That could exist, that does exist, really. NSA, the GCHQ stuff.*

However, thing came less naturally when the artists needed to understand PROV more deeply and create narrative content using the PROV structures. In order for the artists to create provenance graphs and include them in a meaningful way into the narrative, they had to understand both the concept and the PROV standard.

*Dominic: It was a complete mindfuck for me. I don't know about Adam. It just went straight over my head. I was really sitting there thinking: 'I don't know what the fuck they are talking about.' It took me quite a long time to understand what was happening there.*

*Adam: To get your head around it at the start it seemed kind of confusing. And every time it got re-explained it sounded kind of different. But then also you were kind of changing, adapting the model to fit around the game in a way.*

Naturally, the artists would not have needed to pay attention to the PROV standard in any of their other games.

*Adam: In a normal project we would have never looked at any of that because it wouldn't have mattered, it would have been irrelevant. If we were doing a project where you had to click on graphs and choose the bits that were wrong, we would have just made graphs up. We could have fudged it a lot of the time. It wouldn't have to be correct as far as the provenance model goes.*

Eventually, the artists were able to understand provenance well enough to make it an integral part of the narrative. However, combining the provenance-inspired narrative episodes with explicit elements of provenance proved to be more of a challenge.

*Adam: I would have the script written for that episode, and then I would be like: 'Where's an object? Where's an object in this?' It's gotta have a receipt. There has to be a record of something that's happened. A lot of it was people talking and it was really hard sometimes to pick the points where provenance might actually be relevant and not completely forced.*

For developing an episode, they would first write the narrative and then (together with research team) plan what a potential provenance graph for this episode could be.

*Dominic: We'd write the narrative, then we'd probably have a meeting here [at the university], we'd talk about whatever we talked about what needs to be done about it*



*provenance-y research-y wise. Then we go back and pull our hair out trying to find points within the narrative where we can make that work. Then we'd rewrite the narrative in order to make it work. I think for all of us that's where the problems were.*

For them, the narrative was always the starting point – the narrative idea always came first.

*Adam: The provenance wouldn't exist without the narrative. I can't see how you could have the provenance before the narrative. The provenance is the provenance of the narrative.*

Several episodes into the game, the artists became more skilled at including explicit provenance into the narrative from the beginning. This is how they responded when asked whether their writing process was driven by the need to include provenance:

*Adam: That developed. At the start it was just writing the episode. But then towards the end it was more thinking about 'yeah, not paying with cash but paying with credit card'. We deliberately put it in that they were paying with card so that we knew we had a provenance point there whether we used it or not.*

*Dominic: So looking at the narrative and asking: 'What can we track within this, and what can we put in this to track. What was available in 1962 or whenever? But rather looking at that while writing rather than writing, looking at the provenance and then shoehorning provenance into it somehow.*

*Adam: The Dallas one. I was so happy when I discovered that credit cards in America had been issued by that time. They had only been issued the year before, it was the first credit card. I was so pleased. I didn't know how to fix that fucking episode at that time.*

However, having to add these *provenance points* into the narrative was still seen as obstructing.

*Dominic: I suppose the making of the provenance graph is going alongside at this particular point. But the narrative leads it always. But because of the nature of the provenance graph and what's in it, that sort of informs the narrative, and the narrative has to take some sharp right turns instead of wiggly-wiggly, like it usually does. So that was a bit frustrating and difficult to deal with.*

Ultimately, from a player perspective the use of provenance as a key narrative vehicle did not bring up the same frustrations and difficulties as it did for the artists. Participants did not feel that PROV seemed forced onto the narrative and instead accepted it as part of the world they were playing in.

*P3: "At the beginning I was very aware [that this was a research project] because it was very fresh in my mind that I'd signed that consent form that I'll have this data gathered, but as it went on I actually started to forget that this was developed by a group of research students. I was like this is just a fun game that somebody made involving provenance."*

*P2: “Obviously I knew up front [that this was a research project] and took no notice of it the rest of the game.”*

*P8: “It wasn't shoved down my throat throughout the game, which is a good thing, but I was constantly aware that that was a [...] university study.”*

## **Tension 2: Creation of research-relevant content by artists**

Resolving the tension of translating research questions into gameplay elements is naturally the most crucial one for a successful research game. Co-creating such a game together with artists will already improve the game in this regard. However, it is important to consider that the artists first need to be introduced to the underlying research background before they can help transforming it. A more than surface level understanding of the subject matter is hereby crucial as only this way they can create a meaningful game experience that embraces it fully. In the case of AoM, provenance proved to be a rather tricky concept for the artists to understand as for example discussed as part of vignette 2. In order to create a successful game, it is therefore crucial that the researchers help the artists gain an understanding, and make sure that this understanding is in fact correct.

## **Vignette 3: Player-created provenance**

In the ninth episode of the game the Groundsman gave players the task to create a provenance graph. Cr0n had acquired several documents from different sources, and the

players had to make sense of them and report their findings as a provenance graph as well as a written interpretation about their understandings of what had happened.

The documents at their disposal were: a picture of a wooden doll including a brief description, a picture of a spruce tree from a specific district, a photo of someone felling a tree (with date and GPS location), a Tree Felling Request Form (filled in and granted), a Christmas Gifting Request form (filled in and granted), an order as part of the Orphan Rehoming Scheme, and a personal file of Sulim Kwiatkowa. When looking at the broader picture that these documents painted, players came to the (correct) conclusion that the father of a girl called Karina made her a wooden doll as a Christmas present. The same girl was later orphaned and given a new identity by the Ministry of Provenance, which in turn revealed the true past of one of the characters the players had read about in the main narrative.

One of our provenance-related research question was whether people would be able to create their own provenance graphs, and what form these player-created graphs would take on. Adding this as an episode of the main narrative ensured us that all players had to do this in order to progress with the game. We deliberately let players choose the tools they used to create the provenance graph. This was done on the one hand to not influence them in how to represent the different elements, but also because it did not require additional

implementation. We would then be able to compare the outputs by the different players and use their reports to interpret their graphs.

For the artists, this episode posed a big challenge. Similar to the players, it was also a new type of task for them, and they had to figure out which documents to provide to the players that was not only sufficient to understand the events but also engaging.

*Adam: I find it difficult thinking about the graph the other way around, and what they [the players] needed to create it. We had to give them more information than we had in order to be able to create a graph, in a way. 'Cause I've got the background knowledge of all the other graphs, and things, and the stuff that's happened. Whereas the players, we were asking them to create a graph they had never seen before by just giving them certain documents. And I had to create documents that we probably would not have had in the real graph, if we were doing the graph.*

In general, they found it difficult to fit into the narrative, and would certainly not have done so if they had full creative control over the game.

*Dominic: If we had just been making a game, there is no way we would have done it.*

*Adam: I mean also, if you saw a girl 5 years ago with a doll, and you see her 5 years later with the same doll, and she looked 5 years older and she had the same doll, you'd just assume it was the same girl. You wouldn't go and draw a bloody graph about where the*

*tree got cut down and who got permission to give her a doll for Christmas. It was a bit mental in that respect. I know in the world of the Ministry it's normal.*

The episode also proved to be a hurdle for many players. Suddenly, they had to create their own content while before they were just manipulating and inspecting existing content. In the end, several players stopped playing the game instead of engaging with the episode.

*P1: The final reason why I just dropped out was the last report [wooden doll] that I was working on. I couldn't quite figure out what you were looking for with one of the answers. And I just kinda let it go. And the next chapter was announced and every time a new chapter was announced and I would go log in and look but I wasn't able to progress onto it cause I hadn't done that final report. And I thought of writing and saying "can you just help me?" Or writing in character and saying "oh I can't fill out that for some reason, can you put me on to the next chapter?" But by that time I was kind of unhappy with the amount of graphs that I needed to do so those two things kept me from progressing on to the end. So if I hadn't been stopped at that report I probably would have gone on to the next chapters.*

#### **Vignette 4: Provenance of Provenance**

Similar to the task of creating provenance, one of the final episodes of the game was dealing with a new concept: provenance of provenance. Players had to acquire several documents that described how people created and edited other provenance documents. By

looking at these documents and paying close attention to which persons were responsible for which activities, players had to find out who was responsible for suspicious actions, e.g. changing nodes in a top secret classified document which suggested a direct involvement in the MoP conspiracy. Players also had access to four additional provenance graphs describing planned routes of ships important for the final stage of the Ministry's plan. All of these graphs described the same planned activity, but only one of them was actually truthful. Their investigation with regards to the provenance of provenance allowed them to identify the correct graph, which in turn enabled them to relay the final meeting point of the ships to the Groundsman for the final strike against the Ministry.

Like the episode where players had to create their own provenance graphs, the concept of provenance of provenance was brought into the game with a specific research question in mind: Do people understand the concept of provenance of provenance and what do they make of it? We chose the penultimate episode for this task as this would assure that players already had a strong understanding of provenance. Unlike with the creation of provenance, this task did not provide us with any direct data, but was instead, going to be one of the topics in the post-game questionnaires and interviews. To prepare players for this task, we also created daily graphs at the MoP Intranet that dealt with provenance of provenance as an introduction to the concept.

*Adam: The provenance of provenance stuff I personally thought it was a bit shitty, I didn't like it. I found the concept to be hard to understand anyway. [...] I think by that point we had pushed people too far with what we were asking them to do. The more complicated it got, rather than getting them reengaged it started to turn them off at some point. [...] Admittedly, on that last provenance of provenance one I had pretty much given up.*

As with scaffolding narrative to support player-created-provenance, the artists had a hard time integrating a conceptually difficult structure into the narrative. Adding an additional layer of provenance increased the complexity they had to wrap their heads around, which made it even harder for them to fit it into the narrative. In the end, one of the researchers had to create the relevant provenance graphs to make sure they were an accurate representation.

Only a handful of players played the game to the point where they were confronted with provenance of provenance graphs on the Cr0n side. Due to the linear nature of the narrative, only participants who had stuck through with the story could encounter it. There was no way to speed through the first few episodes to ensure that more players encountered provenance of provenance. In addition, due to the complexity of the concept we could not have introduced it any earlier.



Two participants who encountered Provenance of Provenance were amongst those interviewed. One of them commented on its vital role in ensuring system correctness, drawing on his own job as an example.

*P2: Reviewing the reviewer I think is essential, and provenance of provenance is a logical extension of provenance.*

Another saw it purely through the lens of the game and found it very difficult to engage with, and ultimately did not enjoy that aspect:

*P4: They were horrific. Just too many nodes and not enough information.*

### **Tension 3:** Translating research questions into engaging gameplay elements

The last tension we encountered was the desire for the artists to create a game that would live up to their vision, i.e. something that fits their oeuvre. On the one hand, the artists had (almost) full reign over the unfolding narrative, and on the other they were regularly required to add elements to the game they would not normally have (e.g. provenance of provenance). They rightfully emphasized the importance of making the game first and foremost engaging. In order to support this, we supplemented the game with non-research related tasks and content. The most substantial example for this was the 24h hacking challenge (Experience 6) that was solely created to improve player experience without having any direct value for our research agenda. Smaller and more regular examples

included the addition of Cr0n missions that did not feature any provenance graphs but instead consisted of alternative types of riddles. Perhaps unsurprisingly, we received very positive player feedback regarding these elements which in the end helped with player retention. At the same time, we also had to make sure that the artists did in fact focus on the bigger picture of the game. Therefore, we had to intervene when they for example engaged too much with a few hyperactive players instead of trying to improve the engagement of quieter and less active players.

#### **Vignette 5: Direct interaction with players**

Players were able to send in-character messages to the Cr0n headquarters as well as within the MoP intranet to other employees (but not other players).

The communication with NPCs did not serve a specific research purpose; however, it allowed players to seek help and report technical difficulties without breaking character. It also provided us with opportunities to directly ask some of the best performers on the MoP intranet to describe how they went about solving the provenance graphs. We did this to enrich the world (as the interview would be featured in an in-game magazine) and also probe how they were interacting with the provenance graphs. In a way unsurprisingly, the players reacted completely, even overly, in-character, and thus were reluctant to offer much detail in order to not reveal their secrets and appear to be bad employees. The result was

that this communication channel ended up being useful in terms of world building, but entirely unhelpful for collecting real data from participants.

Several players regularly used the communication channels for tips on how to solve the graphs and did so in-character. The artists would then likewise reply in-character trying to put them back on track (and usually succeeding).

The artists enjoyed these direct interactions with the players, as it allowed them to react to the players' thoughts and feelings. They used them to let players explore aspects of the main narrative that had not been covered. This resulted in a situation where several of the more active players were sending the majority of emails, who were then encouraged by any response to make even more use of these communication channels. As a consequence, the artists were spending a significant amount of time engaging with this handful of very active players. This was primarily because they were enjoying it:

*Dominic: The email interaction bits were quite good fun, I'd say, all of it, I mean from all of us that were writing emails. I enjoyed that quite a lot, it was quite amusing. And the system, that was quite easy to deal with.*

However, we saw this as problematic since at that same time there were other players that could have profited from such direct interaction but did not receive it because they were not actively seeking out contact with the NPCs. Thus, whilst the artists were enjoying the depth

of these individual interactions, we were more interested in covering the breadth our participants with at least some level of interaction. To that end, the artists were eventually instructed to engage less with hyperactive players and instead to focus on trying to engage the quieter ones. This did not go unnoticed by our players as players who communicated with NPCs generally enjoyed this level of interaction. They tried to get MoP promotions by play-acting as a double spy, or to uncover more about the past of the Groundsman. They explicitly noted the dip in direct interaction towards the end of the game.

*P1: I thought it was OK. I know that some of the other people involved in the game got really literal about it and started finding like small inconsistencies. Like there was a ministry of provenance worker was emailing some of them and some of her statements were inconsistent so they got very angry about that. [laughs] So I didn't take it quite that seriously. I know other people really liked that aspect of one on one correspondence but they became hostile towards the end that it wasn't proceeding as they wanted or they thought that the people weren't remaining in character as much as they wanted.*

*P4: I love the interaction with the Groundsman. He was great. He was the best thing since sliced bread, literally. He was brilliant. And I don't know who it was, but there was one member of the PM team. He would send emails in game, and I thought he was quite witty.*

*He was quite funny. And I quite enjoyed having a banter with him. That was good. I enjoyed that. But towards the end that stopped.*

### **Vignette 6: 24h Hacking Challenge**

About halfway through the game the players were invited to participate in a hacking attack staged by Cr0n to get access to some of MoP's more secure servers. In order to break through the firewall of the server, players had to decode cryptic messages and enter the correct passphrase. Unlike the rest of the game, all players were working together on this. Each hour, a new code sequence would be released, and in the proceeding 60 minutes the right passphrase had to be entered – otherwise the whole mission would fail. It was sufficient if just one player entered the correct code. Since the narrative relied on the mission succeeding, each passphrase was automatically entered by the system itself a few minutes before the time ran out if required, ensuring the success of the mission.

This was an element of the game that, as with the MoP intranet, we developed as a purely ludic mechanic to make the game more exciting for the players. The riddles were not connected to technical provenance and as such it did not play any role in the research outcomes, only in exploring the AoM fantasy. The challenge proved to be very popular among both the artists and the players.

*Adam: One of my favorite bits was the 24h hacking thing, I have to say. That was one of my favorite bits of the whole thing. Where they had to crack the codes and put them in. I was surprised that the people actually played it.*

The artists also provided reasons for this:

*Dominic: It was more vital, wasn't it? Doing stuff that's live, it just feels more vital.*

*Adam: Even though I wasn't entering codes, I was looking at it, and seeing whether they were getting them. And trying to guess whether the computer itself had really done it. Yeah. And then looking at the thing and seeing someone had. Yeah, I found that really interesting. Even though I wasn't actively participating in it. I found it nice the fact that it was an actual live thing, even though the people weren't all there, they were virtually there.*

For the artists this part of the game was much closer to how they would normally interact with their players: by staging real-time interventions. Usually they would organize an event in the real world so that actors could interact directly with players. Unlike the rest of AoM, this was the only time when the players actually had time pressure for doing something, and where it wasn't a single-player experience.

The feedback by the players was largely positive concerning this event. As it was completely optional it enabled the ones who were interested in it to really invest their day in this challenge, while others could just ignore it.

*P1: And I liked the little one day or 24 hour event where we solved puzzles.*

*P2: When I did the 24 hour code breaking, that was really good. That was REALLY good.*

*It was really time consuming and it was really active and there was a lot of chatter about that.*

*P8: [who had dropped out of the game] So when it was the code breaking challenge, because I didn't have to earn that through doing provenance graphs like before, that's when I got back in, because I didn't have to jump through hoops to get to it.*

Not all players enjoyed it however, and those that did not enjoy it cited its difficulty.

*P6: Like there was this hacking thing where you had to break one of the codes. And I really just couldn't do it and it really put me off cause it was on for ages and you had like 24 hours to break this one and figure out these. I had no idea how to do it and it was absolutely frustrating.*

## Mitigating Tensions

As we have seen in the previous section, the collaboration between domain experts and game designers/developers is not without challenges. Perhaps the most obvious one is the need to translate the domain content under investigation into compelling game mechanics (Winn, 2009). In order to achieve this, it is for example necessary that the game designers

elicit knowledge from the domain experts (van Roessel & van Mastrigt-Ide, 2011) and that the different mind sets (and goals) of game designers and domain experts are consolidated (Cheng et al., 2016). In the existing literature, these challenges have been looked at from the perspective of the game designers. However, in research games such as AoM, the researchers take on the role of domain expert while a third party (i.e. artists) are responsible for the game design. In the case of AoM, as the researchers we also implemented the game design, but it is easy to imagine a case where the researchers are solely responsible for the domain expertise. Looking at the collaboration from this angle allows us to extend the existing knowledge and ultimately provide mitigation strategies to overcome the tensions that we have identified. In the following we will present several ways that could be employed in future collaborations between artists and researchers. Table 1 gives an overview of the relationship between the tensions and our proposed mitigation strategies.

Table 1. Tensions and their mitigation strategies.

<b>Tension</b>	<b>Mitigation Strategy</b>
Tension 1: Translating research questions into engaging gameplay elements	1) Cultivating true co-production 2) Preventing research dependencies



Tension 2: Creation of research-relevant content by artists	3) Proactive transfer of domain knowledge 4) Providing appropriate tools and toolchains
Tension 3: Artistic vision conflicting with research agenda	5) Valuing non-research elements 6) Supporting artistic freedom
Meta	7) Joint responsibility for mitigation strategies

### 1) Cultivating true co-production

A successful research game needs to be an engaging game that at the same time provides enough opportunities to investigate the underlying research questions. In traditional serious games development set-ups, the game designers are commissioned by the domain experts who in turn become clients. Formally, this was also the case with AoM, however, throughout the process we made sure that it was indeed a case of co-production instead of a strictly hierarchical set-up. The importance of such a close relationship on as equal as possible terms was emphasized before: *“Good applied game design [...] happens in the synergy between game designer and domain expert”* (van Roessel & van Mastrigt-Ide, 2011, p9). Over the course of the development process of AoM we had regular meetings with the artist group where the emphasis was not just on progress reports but instead turned into joint design sessions where we openly discussed issues with the game from various perspectives and worked together to overcome these. While this was not a typical set-up for

their work, the artists quickly adapted to this approach. This close involvement increased team identification (Van Der Vegt & Bunderson, 2005), and very directly benefitted our collaboration. We could discuss provenance related questions directly, and it also made sure the project stayed on course throughout preventing the need for sudden and impactful vetoes that might have been necessary with less regular involvement from our side.

## 2) Preventing research dependencies

The Apocalypse of MoP was overall a rather linear experience, due to the expansive and overarching narrative. While players could spend as much time as they liked with the smaller provenance graphs on the MoP intranet, they could only progress in the game by handling the narrative provenance graphs. This was problematic as the graph dealing with provenance of provenance, for example, was only accessible for them once they had reached the final stages of the game. Here, we should have carefully considered before the start of the game how to expose as many players as possible with this type of content. Our advice is to very clearly define the (inter)dependency of the research questions under investigation and planning not just *how* but also *when* they should be integrated into the gameplay. Which research questions can be tested independently of others – and as such can be satisfied by separate elements of gameplay and do not require players to have progressed deep into the game? Game content relevant for specific research questions

should only be *hidden* behind other game content if it is absolutely necessary, for example, to ensure a certain level of understanding by the player.

### 3) Proactive transfer of domain knowledge

A research game needs to incorporate the relevant domain knowledge in the form of appropriate gameplay. Therefore, this knowledge needs to be transferred from the domain experts to the game designers. In previous work, this elicitation of knowledge has been identified as one of the crucial tasks for the designers (van Roessel & van Mastrigt-Ide, 2011). However, based on our experiences, we believe that the onus lies on the researchers much more than on the game designers / artists. Only the researchers are able to ultimately decide whether the content has been integrated correctly and in a meaningful way (regarding the research questions). In our case, the artists were aware that the main goal of AoM was to see how players understand provenance. However, they did not necessarily care if their understanding of provenance was accurate and appropriate. Here, it was our responsibility to introduce the artists to the overarching concept as well as to explain the finer details of it. As described in the vignettes, this was not an easy task, and problems to understand provenance was a big stumbling block throughout the development process. We could have easily decided to leave the artists in a state of less-than-full understanding, but that would have meant meticulously double-checking the provenance graphs they

produced, with any necessary corrections then creating additional work that could have been prevented easily. There is however no best way on how to teach the domain knowledge and in turn make sure that it is understood properly as this very much depends on the topic at hand and the individuals involved.

#### 4) Providing appropriate tools and toolchains

The role of the artists was to create provenance content, for example in order to create the narrative graphs. It was crucial that they were able to do so by themselves as only this way they could experiment with what they produced before it was made live in the actual game. For this we made use of the third-party software, yEd Graph Editor (yWorks, n.d.). This allowed the artists to create their vision of the graph. They could then upload the resulting file to the admin backend of the game where it was translated into the PROV format. The artists could then interact with the graph the same way as players. Giving the artists access to such a tool likewise allowed them to deepen their understanding as they were able to experiment with it and create graphs quickly and efficiently.

By providing these tools we empowered the artists and made the content integration easy enough for them to focus on the design and not on the technical implementation. Providing bespoke authoring tools to enable frictionless participation by non-technologists in the game development process is of course not a specific demand for research games but is

valuable in any game development process that involves stakeholders with different skillsets (Wetzel, Blum, Jurgelionis, & Oppermann, 2012). It is not necessary to develop these tools from scratch, as in our case the appropriation of existing tools can likewise fill the gap with perhaps less effort.

## 5) Valuing non-research elements

We have mentioned previously that the artistic vision was occasionally at odds with the research agenda of the game. It would have been very easy for us to force the artists to only work on parts of the game that had a direct relevance for the research questions we were investigating. However, while elements like the 24h hacking challenge might not have resulted in more data for our research directly, it had a positive effect on the overall appreciation of the game and as such improved player retention and motivation. Player retention is of course an important issue in commercial videogames with game design being one (if not the main) way of preventing player churn (Allart, Levieux, Pierfitte, Guilloux, & Natkin, 2016; Debeauvais, Nardi, Schiano, Ducheneaut, & Yee, 2011). For a research game, high player retention means that more research data is generated per player thus positively affecting the research agenda. In our case, interventions like the 24h hacking challenge did not only prevent players from dropping out, it also made other players who had previously quit the game re-engage upon hearing about it. It is therefore important for

the researchers to realize the value that these on first sight not research-related elements possess.

## 6) Supporting artistic freedom

Closely related to the previous mitigation strategy, we also found it important to allow for a significant amount of leeway for artists to follow their vision. This vision will not always be aligned with the research agenda, but as discussed, this is not necessarily negative.

Granting artists this freedom goes a long way to keeping everybody working on the research game satisfied and fully motivated. A stricter approach would stifle creativity and in turn moot the main reason to involve artists in the first place as they then would not be able to bring in their specific expertise. Compromises need to be made in such a set-up and the involved parties need to be aware and supportive of each other strengths and skills in order to create a compelling joint piece of work (Miller, 1998).

It is difficult to find the right balance between artistic freedom and making sure the research purpose of the game is kept in mind. An example for this were the direct interactions between the artists and specific players of AoM. Here, we had to step in and redirect this engagement away from few, very active players, to a larger number of less active players in order to focus more on the underlying research agenda.

## 7) Joint responsibility for mitigation strategies

Our final and meta advice is to share the burden of coming up with, and perhaps more importantly, implementing the aforementioned mitigation strategies. If both artists and researchers are aware of the potential tensions that could arise, both parties can focus on how to jointly overcome these. In the case of AoM, we had several team meetings where we openly discussed the needs of the project from a research perspective. At the same time, the group of artists was likewise honest in raising concerns when they felt the need for it. One example was the necessity to help them understand the concept of provenance so that they could weave it into the game narrative. By both of us being aware that this had the potential to turn into a bigger complication along the road, we were able to mitigate the problem by working together on it.

## Conclusion

In this article, we discussed the game Apocalypse of MoP, a research game about provenance, that we co-created together with a group of artists. While the game was successful in providing us with insight into how provenance is perceived, we also had to deal with various tensions arising from our collaboration. We do believe that the involvement of professional game creators outweighs any of the arising tensions – if these are carefully taken into consideration. For us, developing the Apocalypse of MoP in such

an environment was a learning experience. We have gained first-hand insight into what works well and what does not and hope our reflections will help shape and inform future collaborations of this. In order to do so, we have chosen the form of a post-mortem, informed by our own reflections, thoughts and impressions by the involved artists, and feedback from players. We have identified three tensions that recurred throughout the process as evidenced by a collection of six richly described vignettes. The tensions are: 1) Translating research questions into engaging gameplay elements; 2) Creation of research-relevant content by artists; and 3) Artistic vision conflicting with research agenda.

From our experiences we are confident that these tensions can be overcome. We provide seven mitigation strategies in order to overcome the identified tensions: emphasis on co-production; considerations on which parts of a game relate to which research questions; making sure the artists understand and can apply the necessary domain knowledge; providing them with tools to create their own content easily; appreciating the value of non-research related content, and supporting artistic freedom. As there will always be tensions between the research agenda and the artistic vision, it is furthermore up to both the researchers and artists to co-produce mitigation strategies. This is necessary to achieve the desired result of a successful research game - meaning it is both engaging as well as informing the research.



Lastly, our work can also be placed in the research context of serious games where multidisciplinary teams of domain experts and game developers are the rule rather than the exception. Here, we provide unique perspectives from the former group instead of solely focusing on the latter. We believe understanding all groups involved in such a collaborative game creation process is highly valuable for practitioners in this field.

## Acknowledgements

We would like to thank Urban Angel who made the game possible, especially Adam Sporne and Dominic Shaw. Thanks also to Stuart Moran for his support in running player interviews. This research was approved by the research ethics committee of the School of Computer Science at the University of Nottingham.

## Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

This work described in this article is supported by the Engineering and Physical Sciences Research Council grant number EP/I011587/1.

## References

- Allart, T., Levieux, G., Pierfitte, M., Guilloux, A., & Natkin, S. (2016). Design influence on player retention: A method based on time varying survival analysis. In *Proceedings of the 2016 IEEE Conference on Computational Intelligence and Games (CIG)* (pp. 1–8). doi:10.1109/CIG.2016.7860421
- Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., Freitas, S. de, Louchart, S., ... Gloria, A. D. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*, 46(2), 391–411. doi:10.1111/bjet.12113
- Bachour, K., Wetzel, R., Flintham, M., Huynh, T. D., Rodden, T., & Moreau, L. (2015). Provenance for the people: An HCI perspective on the W3C PROV standard through an online game. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2437–2446). New York, NY: ACM. doi:10.1145/2702123.2702455

- Barthel, R., Leder Mackley, K., Hudson-Smith, A., Karpovich, A., Jode, M., & Speed, C. (2013). An internet of old things as an augmented memory system. *Personal Ubiquitous Computing* 17(2), 321–333. doi:10.1007/s00779-011-0496-8
- Beaubien, J. M., & Baker, D. P. (2004). The use of simulation for training teamwork skills in health care: how low can you go? *BMJ Quality & Safety*, 13(suppl 1), i51–i56. doi:10.1136/qshc.2004.009845
- Benford, S., Crabtree, A., Flinham, M., Drozd, A., Anastasi, R., Paxton, M., ... Row-Farr, J. (2006). Can you see me now? *ACM Transactions on Computer-Human Interaction*, 13(1), 100–133. doi:10.1145/1143518.1143522
- Benford, S., Greenhalgh, C., Crabtree, A., Flinham, M., Walker, B., Marshall, J., ... Row-Farr, J. (2013). Performance-led research in the wild. *ACM Transactions on Computer-Human Interaction*, 20(3), 14:1–14:22. doi:10.1145/2491500.2491502
- Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., & Rodden, T. (2012). Uncomfortable interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2005–2014). New York, NY: ACM. doi:10.1145/2207676.2208347
- Benford, S., Hazzard, A., Chamberlain, A., Glover, K., Greenhalgh, C., Xu, L., ... Darzentas, D. (2016). Accountable artefacts: The case of the Carolan Guitar. In

*Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1163–1175). New York, NY, USA: ACM. doi:10.1145/2858036.2858306

Catalano, C. E., Luccini, A. M., & Mortara, M. (2014). Guidelines for an effective design of serious games. *International Journal of Serious Games*, 1(1). doi:10.17083/ijsg.v1i1.8

Cheng, J., Putnam, C., & Guo, J. (2016). Always a tall order: Values and practices of professional game designers of serious games for health. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 217–228). New York, NY, USA: ACM. doi:10.1145/2967934.2968081

Davis, S. B., Moar, M., Jacobs, R., Watkins, M., Riddoch, C., & Cooke, K. (2006). ‘Ere be dragons: Heartfelt gaming. *Digital Creativity*, 17(3), 157–162.

De Nies, T., Magliacane, S., Verborgh, R., Coppens, S., Groth, P., Mannens, E., & Van de Walle, R. (2013). Git2PROV: Exposing version control system content as W3C PROV. In *Poster and Demo Proceedings of the 12th International Semantic Web Conference* (Vol. 1035, pp. 125–128). Retrieved from [http://ceur-ws.org/Vol-1035/iswc2013\\_demo\\_32.pdf](http://ceur-ws.org/Vol-1035/iswc2013_demo_32.pdf)

Debeauvais, T., Nardi, B., Schiano, D. J., Ducheneaut, N., & Yee, N. (2011). If you build it they might stay: Retention mechanisms in World of Warcraft. In *Proceedings of the 6th International Conference on Foundations of Digital Games* (pp. 180–187). New York, NY, USA: ACM. doi:10.1145/2159365.2159390

Del Rio, N., & Da Silva, P. P. (2007). Probe-it!: Visualization support for provenance. In *Proceedings of the 3rd International Conference on Advances in Visual Computing - Volume Part II* (pp. 732–741). Berlin, Heidelberg: Springer-Verlag. Retrieved from <http://dl.acm.org/citation.cfm?id=1779090.1779169>

Evans, E., Flintham, M., & Martindale, S. (2014). The Malthusian Paradox: performance in an alternate reality game. *Personal and Ubiquitous Computing*, 18(7), 1567–1582.  
doi:10.1007/s00779-014-0762-7

Fischer, J. E., Jiang, W., & Moran, S. (2012). AtomicOrchid: a mixed reality game to investigate coordination in disaster response. In *Proceedings of the 11th international conference on Entertainment Computing* (pp. 572–577). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-33542-6\_75

Ghorashi, S., & Jensen, C. (2012). Leyline: Provenance-based search using a graphical sketchpad. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (pp. 2:1–2:10). New York, NY, USA: ACM.  
doi:10.1145/2391224.2391226

Groth, P., & Moreau, L. (Eds.). (2013). *PROV-Overview : an overview of the PROV family of documents*. W3C Working Group Note NOTE-prov-overview-20130430, World Wide Web Consortium. Retrieved from <http://travesia.mcu.es/portalnjb/jspui/handle/10421/7487>

Haferkamp, N., Kraemer, N. C., Linehan, C., & Schembri, M. (2011). Training disaster communication by means of serious games in virtual environments. *Entertainment Computing*, 2(2), 81–88. doi:10.1016/j.entcom.2010.12.009

Hart, J., Iacovides, I., Adams, A., Oliveira, M., & Margoudi, M. (2017). Understanding engagement within the context of a safety critical game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (pp. 253–264). New York, NY, USA: ACM. doi:10.1145/3116595.3116633

Herschel, M., Diestelkämper, R., & Lahmar, H. B. (2017). A survey on provenance: What for? What form? What from? *The VLDB Journal*, 26(6), 881–906. doi:10.1007/s00778-017-0486-1

Herschel, M., & Hlawatsch, M. (2016). Provenance: On and behind the screens. In *Proceedings of the 2016 International Conference on Management of Data* (pp. 2213–2217). New York, NY, USA: ACM. doi:10.1145/2882903.2912568

Khaled, R., & Ingram, G. (2012). Tales from the front lines of a large-scale serious game project. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 69–78). New York, NY, USA: ACM. doi:10.1145/2207676.2207688

Light, A. (2014). The allure of provenance: Tracing food through user-generated production information. In *Eat, Cook, Grow: Mixing Human-Computer Interactions with Human-Food Interactions* (pp. 213–226). MIT Press.

Marfisi-Schottman, I., George, S., & Tarpin-Bernard, F. (2010). Tools and methods for efficiently designing serious games. In *Proceedings of ECGBL 2010 The 4th European Conference on Games Based Learning*. Danish School of Education Aarhus University, Copenhagen.

McGarry, G., Tolmie, P., Benford, S., Greenhalgh, C., & Chamberlain, A. (2017). They're all going out to something weird: Workflow, legacy and metadata in the music production process. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 995–1008). New York, NY, USA: ACM.

doi:10.1145/2998181.2998325

Miller, P. (1998). The engineer as catalyst: Billy Kluver on working with artists. *IEEE Spectrum*, 35(7), 20–29. doi:10.1109/6.694345

Missier, P., & Chen, Z. (2013). Extracting PROV provenance traces from Wikipedia history pages. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops* (pp. 327–330). New York, NY, USA: ACM. doi:10.1145/2457317.2457375

Moreau, L. (2010). The foundations for provenance on the web. *Found. Trends Web Sci.*, 2(2–3), 99–241. doi:10.1561/18000000010

Moreau, L., & Missier, P. (Eds.). (2013). *PROV-DM: The PROV data model*. W3C Recommendation REC-prov-dm-20130430, World Wide Web Consortium. Retrieved from <http://www.w3.org/TR/2013/REC-prov-dm-20130430/>

Rubin, V., Conroy, N. J., & Chen, Y. (2015). Towards news verification: Deception detection methods for news discourse. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS48) Symposium on Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium*. doi:DOI: 10.13140/2.1.4822.8166

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Shirinian, A. (2011). Dissecting the postmortem: lessons learned from two years of game development self-reportage. Retrieved July 31, 2018, from [https://www.gamasutra.com/view/feature/134679/dissecting\\_the\\_postmortem\\_lessons\\_.php](https://www.gamasutra.com/view/feature/134679/dissecting_the_postmortem_lessons_.php)

Van Der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management Journal*, 48(3), 532–547. doi:10.5465/amj.2005.17407918



van Roessel, L., & van Mastrigt-Ide, J. (2011). Collaboration and team composition in applied game creation processes. In *Proceedings of the 2011 DiGRA International Conference: Think Design Play*. DiGRA/Utrecht School of the Arts.

von der Pütten, A. M., Klatt, J., Ten Broeke, S., McCall, R., Krämer, N. C., Wetzel, R., ...

Klatt, J. (2012). Subjective and behavioral presence measurement and interactivity in the collaborative augmented reality game TimeWarp. *Interacting with Computers*, 24(4), 317–325. doi:10.1016/j.intcom.2012.03.004

Westera, W., Nadolski, R. J., Hummel, H. G. K., & Wopereis, I. G. J. H. (2008). Serious games for higher education: a framework for reducing design complexity. *Journal of Computer Assisted Learning*, 24(5), 420–432. doi:10.1111/j.1365-2729.2008.00279.x

Wetzel, R., Blum, L., Jurgelionis, A., & Oppermann, L. (2012). Shapes, Marbles and Pebbles: template-based content creation for location-based games. In *Proceedings of IADIS International Conference Game and Entertainment Technologies GET2012*. Lisbon, Portugal.

Wetzel R., Waern A., Jonsson S., Lindt I., Ljungstrand P., Åkesson KP. (2009) Boxed pervasive games: An experience with user-created pervasive games. *Lecture Notes in Computer Science Pervasive Computing*, 220-237. doi:10.1007/978-3-642-01516-8\_16

Winn, B. M. (2009). The design, play, and experience framework. *Handbook of Research on Effective Electronic Gaming in Education*, 1010–1024. doi:10.4018/978-1-59904-808-6.ch058

yWorks. (n.d.). *yEd Graph Editor*. Retrieved from <http://www.yworks.com/yed>

## Author Biographies

**Richard Wetzel** is an Assistant Professor in the School of Information Technology at the Lucerne University of Applied Sciences and Arts. He is co-lead of the research group Immersive Realities and founding member of the Ludic Lab Lucerne. Richard holds a PhD from the Mixed Reality Lab of the University of Nottingham. He investigates mixed reality games and experiences and has developed a deck of ideation cards to support multidisciplinary collaboration between different stakeholders in the design process.

**Khaled Bachour** is a Senior Lecturer in Games Computing and HCI in the School of Computer Science at the University of Lincoln. He holds a PhD in Computer Science from the Swiss Federal Institute of Technology in Lausanne, Switzerland (EPFL) for a dissertation on interactive technologies for Computer-Supported Collaborative Learning. His research focus is on using games as a methodology for HCI research having explored a

wide range of topics from human-agent collaboration to the usability of digital provenance through the medium of research games.

**Martin Flintham** is an Assistant Professor in Computer Science at the University of Nottingham, where he is a member of the Mixed Reality Lab and Horizon Digital Economy Research. His research interests include building novel technology probes and interactive experiences, including games, mixed reality, and the Internet of Things. He has led successful “in the wild” collaborations with the creative industries, leading to several highly cited publications at ACM CHI, and winning the Prix Ars Electronica and two BAFTA nominations.